



## Digital signal processing for fiber nonlinearities [Invited]

**Cartledge, John C.; Guiomar, Fernando P.; Kschischang, Frank R.; Liga, Gabriele; Yankov, Metodi Plamenov**

*Published in:*  
Optics Express

*Link to article, DOI:*  
[10.1364/OE.25.001916](https://doi.org/10.1364/OE.25.001916)

*Publication date:*  
2017

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Cartledge, J. C., Guiomar, F. P., Kschischang, F. R., Liga, G., & Yankov, M. P. (2017). Digital signal processing for fiber nonlinearities [Invited]. *Optics Express*, 25(3), 1916-1936. <https://doi.org/10.1364/OE.25.001916>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Digital signal processing for fiber nonlinearities [Invited]

JOHN C. CARTLEDGE,<sup>1,\*</sup> FERNANDO P. GUIOMAR,<sup>2,6</sup> FRANK R. KSCHISCHANG,<sup>3,7</sup> GABRIELE LIGA,<sup>4,8</sup> AND METODI P. YANKOV<sup>5,9</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, Queen's University, Kingston, ON K7L 3N6, Canada

<sup>2</sup>Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino, Italy

<sup>3</sup>The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada

<sup>4</sup>Optical Networks Group, Department of Electronic and Electrical Engineering, University College London, Torrington Place, London WC1E 7JE, UK

<sup>5</sup>Department of Photonics Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

<sup>6</sup>fernando.guiomar@polito.it

<sup>7</sup>frank@ece.utoronto.ca

<sup>8</sup>gabriele.liga.11@ucl.ac.uk

<sup>9</sup>meya@fotonik.dtu.dk

\*john.cartledge@queensu.ca

**Abstract:** This paper reviews digital signal processing techniques that compensate, mitigate, and exploit fiber nonlinearities in coherent optical fiber transmission systems.

© 2017 Optical Society of America

**OCIS codes:** (060.2330) Coherent communications; (060.2360) Fiber optics communications; (060.4370) Fiber optics links and subsystems; (060.4370) Nonlinear optics, fibers.

## References and links

1. A. Mecozzi and R.-J. Essiambre, "Nonlinear Shannon limit in pseudolinear coherent systems," *J. Lightw. Technol.* **30**(12), 2011–2024 (2012).
2. A. Mecozzi, C. B. Clausen, and M. Shtaif, "Analysis of intrachannel nonlinear effects in highly dispersed optical pulse transmission," *IEEE Photon. Technol. Lett.* **12**(4), 392–394 (2000).
3. L. Dou, Z. Tao, L. Li, W. Yan, T. Tanimura, T. Hoshida, and J. C. Rasmussen, "A low complexity pre-distortion method for intra-channel nonlinearity," in *Optical Fiber Communication Conference* (2011), paper OThF5.
4. Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Multiplier-free intrachannel nonlinearity compensating algorithm operating at symbol rate," *J. Lightw. Technol.* **29**(17), 2570–2576 (2011).
5. T. Oyama, H. Nakashima, S. Oda, T. Yamauchi, Z. Tao, T. Hoshida, and J. C. Rasmussen, "Robust and efficient receiver-side compensation method for intra-channel nonlinear effects," in *Optical Fiber Communication Conference* (2014), paper Tu3A.3.
6. A. Ghazisaeidi, I. Fernandez de Jauregui Ruiz, L. Schmalen, P. Tran, P. Brindel, C. Simonneau, E. Awwad, B. Uscumlic, P. Brindel, and G. Charlet, "Submarine transmission systems using digital nonlinear compensation and adaptive rate forward error correction," *IEEE/OSA J. Lightw. Technol.* **34**(8), 1886–1895 (2016).
7. Y. Fan, L. Dou, Z. Tao, L. Li, S. Oda, T. Hoshida, and J. C. Rasmussen, "Modulation format dependent phase noise caused by intra-channel nonlinearity," in *European Conference on Optical Communication* (2012), paper We.2.C.3.
8. X. Wei, "Power-weighted dispersion distribution function for characterizing nonlinear properties of long-haul optical transmission links," *Opt. Lett.* **31**(17), 2544–2546 (2006).
9. Y. Zhao, L. Dou, Z. Tao, M. Yan, S. Oda, T. Tanimura, T. Hoshida, and J. C. Rasmussen, "Improved analytical model for intra-channel nonlinear distortion by relaxing the lossless assumption," in *European Conference on Optical Communication* (2013), paper P.4.15.
10. Z. Tao, Y. Zhao, W. Fan, L. Dou, T. Hoshida, and J. C. Rasmussen, "Analytical intrachannel nonlinear models to predict the nonlinear noise waveform," *IEEE/OSA J. Lightw. Technol.* **33**(10), 2111–2119 (2015).
11. A. Ghazisaeidi and R.-J. Essiambre, "Calculation of coefficients of perturbative nonlinear pre-compensation for Nyquist pulses," in *European Conference on Optical Communication* (2014), paper We.1.3.3.
12. Y. Zhao, L. Dou, Z. Tao, Y. Xu, T. Hoshida, and J. C. Rasmussen, "Nonlinear noise waveform estimation for arbitrary signal based on Nyquist nonlinear model," in *European Conference on Optical Communication* (2014), paper P.5.8.
13. T. Oyama, H. Nakashima, T. Hoshida, T. Tanimura, Y. Akiyama, Z. Tao, and J. C. Rasmussen, "Complexity reduction of perturbation-based nonlinear compensator by sub-band processing," in *Optical Fiber Communication Conference* (2015), paper Th3D.7.

14. P. Poggiolini, A. Nespola, Y. Jiang, G. Bosco, A. Carena, L. Bertignono, S. M. Bilal, S. Abrate, and F. Forghieri, "Analytical and experimental results on system maximum reach increase through symbol rate optimization," *J. Lightw. Technol.* **34**(8), 1872–1885 (2016).
15. Q. Zhuge, M. Reimer, A. Borowiec, M. O'Sullivan, and D. V. Plant, "Aggressive quantization on perturbation coefficients for nonlinear pre-distortion," in *Optical Fiber Communication Conference* (2014), paper Th4D.7.
16. Y. Gao, J. C. Cartledge, A. S. Karar, and S. S.-H. Yam, "Reducing the complexity of perturbation based nonlinearity pre-compensation using symmetric EDC and pulse shaping," *Opt. Express* **22**(2), 1209–1219 (2014).
17. Z. Li, W.-R. Peng, F. Zhu, and Y. Bai, "MMSE-based optimization of perturbation coefficients quantization for fiber nonlinearity," *IEEE/OSA J. Lightw. Technol.* **33**(20), 4311–4317 (2015).
18. M. Malekiha and D. V. Plant, "Adaptive optimization of quantized perturbation coefficients for fiber nonlinearity compensation," *IEEE Photon. J.* **8**(3), 7200207 (2016).
19. E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *IEEE/OSA J. Lightw. Technol.* **26**(20), 3416–3425 (2008).
20. E. Ip, "Nonlinear compensation using backpropagation for polarization-multiplexed transmission," *IEEE/OSA J. Lightw. Technol.* **28**(6), 939–951 (2010).
21. X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, and G. Li, "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing," *Opt. Express* **16**(2), 880–888 (2008).
22. D. Rafique and A. D. Ellis, "Impact of signal-ASE four-wave mixing on the effectiveness of digital back-propagation in 112 Gb/s PM-QPSK systems" *Opt. Express* **19**(4), 3449–3454 (2011).
23. G. Gao, X. Chen, and W. Shieh, "Influence of PMD on fiber nonlinearity compensation using digital back propagation," *Opt. Express* **20**(13), 14406–14418 (2012).
24. G. Liga, T. Xu, A. Alvarado, R. I. Killey, and P. Bayvel, "On the performance of multichannel digital backpropagation in high-capacity long-haul optical transmission," *Opt. Express* **22**(24), 30053–30062 (2014).
25. G. Liga, C. Czegledi, T. Xu, E. Agrell, R. I. Killey, and P. Bayvel, "Ultra-wideband nonlinearity compensation performance in the presence of PMD," in *European Conference on Optical Communication* (2016), paper P1.SC3.9.
26. E. Mateo, L. Zhu, and G. Li, "Impact of XPM and FWM on the digital implementation of impairment compensation for WDM transmission using backward propagation," *Opt. Express* **16**(20), 16124–16137 (2008).
27. R. Dar and P. Winzer, "On the limits of digital back-propagation in fully loaded WDM systems," *IEEE Photon. Technol. Lett.* **28**(11), 1253–1256 (2016).
28. A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Express* **22**(13), 16335–16362 (2014).
29. P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "A simple and effective closed-form GN model correction formula accounting for signal non-Gaussian distribution," *Opt. Express* **33**(2), 459–473 (2015).
30. M. Secondini, E. Forestieri, and G. Prati, "Achievable information rate in nonlinear WDM fiber-optic systems with arbitrary modulation formats and dispersion maps," *IEEE/OSA J. Lightw. Technol.* **31**(23), 3839–3852 (2013).
31. D. Lavery, D. Ives, G. Liga, A. Alvarado, S. J. Savory, and P. Bayvel, "The benefit of split nonlinearity compensation for single channel optical fiber communications," *IEEE Photon. Technol. Lett.* **28**(17), 1803–1806 (2016).
32. A. D. Ellis, M. E. McCarthy, M. a. Z. Al-Khateeb, and S. Sygletos, "Capacity limits of systems employing multiple optical phase conjugators," *Opt. Express* **23**(16), 20381–20393 (2015).
33. P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *IEEE/OSA J. Lightw. Technol.* **32**(4), 694–721 (2014).
34. R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Express* **21**(22), 25685–25699 (2013).
35. G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "Suppression of spurious tones induced by the split-step method in fiber systems simulation," *IEEE Photon. Technol. Lett.* **12**(5), 489–491 (2000).
36. M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems* (John Wiley & Sons, 1980).
37. K. V. Peddanarappagari and M. Brandt-Pearce, "Volterra series transfer function of single-mode fibers," *IEEE/OSA J. Lightw. Technol.* **15**(12), 2232–2241 (1997).
38. M. Nazarathy, J. Khurgin, R. Weidenfeld, Y. Meiman, P. Cho, R. Noe, I. Shpanzter, and V. Karagodsky, "Phased-array cancellation of nonlinear FWM in coherent OFDM dispersive multi-span links," *Opt. Express* **16**(20), 15777–15810 (2008).
39. F. P. Guiomar, J. D. Reis, A. L. Teixeira, and A. N. Pinto, "Volterra series transfer function of single-mode fibers," *IEEE Photon. Technol. Lett.* **23**(19), 1412–1414 (2011).
40. F. P. Guiomar, J. D. Reis, A. L. Teixeira, and A. N. Pinto, "Mitigation of intra-channel nonlinearities using a frequency-domain Volterra series equalizer," *IEEE Photon. Technol. Lett.* **20**(2), 1360–1369 (2012).
41. L. Liu, L. Li, Y. Huang, K. Cui, Q. Xiong, F. N. Hauske, C. Xie, and Y. Cai, "Intrachannel nonlinearity compensation by inverse Volterra series transfer function," *IEEE/OSA J. Lightw. Technol.* **30**(3), 310–316 (2012).
42. A. Vannucci, P. Serena, and A. Bononi, "The RP method: a new tool for the iterative solution of the nonlinear Schrödinger equation," *IEEE/OSA J. Lightw. Technol.* **20**(7), 1102–1112 (2002).
43. G. Shulkind and M. Nazarathy, "Nonlinear digital back propagation compensator for coherent optical OFDM based on factorizing the Volterra series transfer function," *Opt. Express* **21**(11), 13145–13161 (2013).
44. A. Bakhshali, W. Y. Chan, Y. Gao, J. C. Cartledge, M. O'Sullivan, C. Laperle, A. Borowiec, and K. Roberts,

- “Complexity reduction of frequency-domain Volterra-based nonlinearity post-compensation using symmetric electronic dispersion compensation,” in *European Conference on Optical Communication* (2014), paper P.3.9.
45. A. Bakhshali, W. Y. Chan, J. C. Cartledge, M. O’Sullivan, C. Laperle, A. Borowiec, and K. Roberts, “Frequency-domain Volterra-based equalization structures for efficient mitigation of intrachannel Kerr nonlinearities,” *IEEE/OSA J. Lightw. Technol.* **34**(8), 1770–1777 (2016).
  46. F. P. Guiomar and A. N. Pinto, “Simplified Volterra series nonlinear equalizer for polarization-multiplexed coherent optical systems,” *IEEE/OSA J. Lightw. Technol.* **31**(23), 3879–3891 (2013).
  47. F. P. Guiomar, S. B. Amado, A. Carena, G. Bosco, A. Nespola, A. L. Teixeira, and A. N. Pinto, “Fully-blind linear and nonlinear equalization for 100G PM-64QAM optical systems,” *IEEE/OSA J. Lightw. Technol.* **33**(7), 1265–1274 (2015).
  48. F. P. Guiomar, S. B. Amado, C. S. Martins, and A. N. Pinto, “Time domain Volterra-based digital backpropagation for coherent optical systems,” *IEEE/OSA J. Lightw. Technol.* **33**(15), 3170–3181 (2015).
  49. S. B. Amado, F. P. Guiomar, N. J. Muga, R. M. Ferreira, J. D. Reis, S. M. Rossi, A. Chiuchiarelli, J. R. F. Oliveira, A. L. Teixeira, and A. N. Pinto, “Low complexity advanced DBP algorithms for ultra-long-haul 400G transmission systems,” *IEEE/OSA J. Lightw. Technol.* **34**(8), 1793–1799 (2016).
  50. F. P. Guiomar, S. B. Amado, C. S. Martins, and A. N. Pinto, “Parallel split-step method for digital backpropagation,” in *Optical Fiber Communication Conference* (2015), paper Th2A.28.
  51. G. Shulkind and M. Nazarathy, “Estimating the Volterra series transfer function over coherent optical OFDM for efficient monitoring of the fiber channel nonlinearity,” *Opt. Express* **20**(27), 29035–29062 (2012).
  52. T. Freckmann, R. Essiambre, P. J. Winzer, G. J. Foschini, and G. Kramer, “Fiber capacity limits with optimized ring constellations,” *IEEE Photon. Technol. Lett.* **21**(20), 1496–1498 (2009).
  53. T. H. Lotz, X. Liu, S. Chandrasekhar, P. J. Winzer, H. Haunstein, S. Randel, S. Corteselli, B. Zhu, and D. W. Peckham, “Coded PDM-OFDM transmission with shaped 256-iterative-polar-modulation achieving 11.15-b/s/Hz intrachannel spectral efficiency and 800-km reach,” *IEEE/OSA J. Lightw. Technol.* **31**(4), 538–545 (2013).
  54. D. S. Millar, T. Koike-Akino, S. Ö. Arik, K. Kojima, K. Parsons, T. Yoshida, and T. Sugihara, “High-dimensional modulation for coherent optical communications systems,” *Opt. Express* **22**(7), 8798–8812 (2014).
  55. A. D. Shiner, M. Reimer, A. Borowiec, S. Oveis Gharan, J. Gaudette, P. Mehta, D. Charlton, K. Roberts, and M. O’Sullivan, “Demonstration of an 8-dimensional modulation format with reduced inter-channel nonlinearities in a polarization multiplexed coherent system,” *Opt. Express* **22**(17), 20366–20374 (2014).
  56. R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, “On shaping gain in the nonlinear fiber-optic channel,” in *International Symposium on Information Theory*, 2794–2798 (2014).
  57. B. P. Smith and F. R. Kschischang, “A pragmatic coded modulation scheme for high-spectral-efficiency fiber-optic communications,” *J. Lightw. Technol.* **30**(13), 1–7 (2012).
  58. L. Beygi, E. Agrell, J. M. Kahn, and M. Karlsson, “Rate-adaptive coded modulation for fiber-optic communications,” *J. Lightw. Technol.* **32**(2), 333–343 (2014).
  59. T. Fehenberger, G. Böcherer, A. Alvarado, and N. Hanik, “LDPC coded modulation with probabilistic shaping for optical fiber systems,” in *Optical Fiber Communication Conference* (2015), paper Th2A.23.
  60. F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idler, “Rate adaptation and reach increase by probabilistically shaped 64-QAM: an experimental demonstration,” *J. Lightw. Technol.* **34**(7), 1599–1609 (2016).
  61. A. Ghazisaeidi, I. D. J. Ruiz, R. Rios-Müller, L. Schmalen, P. Tran, P. Brindel, A. C. Meseguer, Q. Hu, F. Buchali, G. Charlet, and J. Renaudier, “65 Tb/s transoceanic transmission using probabilistically-shaped PDM-64QAM,” in *European Conference on Optical Communication* (2016), paper Th.3.C.4.
  62. M. P. Yankov, D. Zibar, K. J. Larsen, L. P. B. Christensen, and S. Forchhammer, “Constellation shaping for fiber-optic channels with QAM and high spectral efficiency,” *IEEE Photon. Technol. Lett.* **26**(23), 2407–2410 (2014).
  63. M. P. Yankov, F. Da Ros, E. P. da Silva, S. Forchhammer, K. J. Larsen, L. Oxenløwe, M. Galili, and D. Zibar, “Constellation shaping for WDM systems using 256QAM/1024QAM with probabilistic optimization,” *J. Lightw. Technol.* **34**(22), 5146–5156 (2016).
  64. “2006 Steele Prizes,” *Notices of the AMS* **53**(4), 464–470 (2006).
  65. A. Hasegawa and F. Tappert, “Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers I. Anomalous dispersion,” *App. Phys. Lett.* **23**(3), 142–144 (1973).
  66. M. I. Yousefi and F. R. Kschischang, “Information transmission using the nonlinear Fourier transform, Part I: Mathematical tools,” *IEEE Trans. Inform. Theory* **60**(7), 4312–4328 (2014).
  67. —, “Information transmission using the nonlinear Fourier transform, Part II: Numerical methods,” *IEEE Trans. Inform. Theory* **60**(7), 4329–4345 (2014).
  68. —, “Information transmission using the nonlinear Fourier transform, Part III: Spectrum modulation,” *IEEE Trans. Inform. Theory* **60**(7), 4346–4369 (2014).
  69. V. E. Zakharov and A. B. Shabat, “Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media,” *Soviet J. of Exp. and Theo. Phys.* **34**(1), 62–69 (1972).
  70. E. Meron, M. Feder, and M. Shtaif, “On the achievable communication rates of generalized soliton transmission systems,” *arXiv:1207.0297v2* (2012).
  71. J. E. Prilepsky, S. A. Derevyanko, and S. K. Turitsyn, “Nonlinear spectral management: Linearization of the lossless fiber channel,” *Opt. Express* **21**(20), 344–367 (2013).
  72. J. E. Prilepsky, S. A. Derevyanko, K. J. Blow, I. Gabitov, and S. K. Turitsyn, “Nonlinear inverse synthesis and

- eigenvalue division multiplexing in optical fiber channels,” *Phys. Rev. Lett.* **113**(1), 013901 (2014).
73. S. Hari, F. Kschischang, and M. Yousefi, “Multi-eigenvalue communication via the nonlinear Fourier transform,” in *Biennial Symposium on Communications* (2014), pp. 92–95.
  74. I. Tavakkolnia and M. Safari, “Signalling over nonlinear fibre-optic channels by utilizing both solitonic and radiative spectra,” in *European Conference on Networks and Communications* (2015), pp. 103–107.
  75. H. Bülow, “Experimental demonstration of optical signal detection using nonlinear Fourier transform,” *J. Lightw. Technol.* **33**(7), 1433–1439 (2015).
  76. Z. Dong, S. Hari, T. Gui, K. Zhong, M. Yousefi, C. Lu, P.-K. Alexander Wai, F. Kschischang, and A. Lau, “Nonlinear frequency division multiplexed transmissions based on NFT,” *IEEE Photon. Technol. Lett.* **27**(15), 1621–1623 (2015).
  77. H. Terauchi, Y. Matsuda, A. Toyota, and A. Maruta, “Noise tolerance of eigenvalue modulated optical transmission system based on digital coherent technology,” in *OptoElectronics and Communication Conference and Australian Conference on Optical Fibre Technology* (2014), pp. 778–780.
  78. V. Aref, H. Bülow, K. Schuh, and W. Idler, “Experimental demonstration of nonlinear frequency division multiplexed transmission,” in *European Conference on Optical Communication* (2015), paper Tu.1.1.2.
  79. K. Schuh, V. Aref, H. Bülow, and W. Idler, “Collision of QPSK modulated solitons,” in *Optical Fiber Communication Conference* (2016), paper W2A.33.
  80. H. Bülow, V. Aref, K. Schuh, and W. Idler, “Experimental nonlinear frequency domain equalization of QPSK modulated 2-eigenvalue soliton,” in *Optical Fiber Communication Conference* (2016), paper Tu2A.3.
  81. V. Aref, H. Bülow, and K. Schuh, “On spectral phase estimation of noisy solitonic transmission,” in *Optical Fiber Communication Conference* (2016), paper W3A.3.
  82. S. Wahls and H. V. Poor, “Introducing the fast nonlinear Fourier transform,” in *IEEE International Conference on Acoustics, Speech and Signal Processing* (2013), pp. 5780–5784.
  83. S. Wahls and H. Poor, “Fast inverse nonlinear Fourier transform for generating multi-solitons in optical fiber,” in *IEEE International Symposium on Information Theory* (2015), pp. 1676–1680.
  84. S. Civelli, L. Barletti, and M. Secondini, “Numerical methods for the inverse nonlinear Fourier transform,” in *Tyrrhenian International Workshop on Digital Communications* (2015), pp. 13–16.
  85. Q. Zhang and T. Chan, “A Gaussian noise model of spectral amplitudes in soliton communication systems,” in *IEEE International Workshop on Signal Processing Advances in Wireless Communications* (2015), pp. 455–459.
  86. —, “A spectral domain noise model for optical fibre channels,” in *IEEE International Symposium on Information Theory* (2015), pp. 1660–1664.
  87. N. Shevchenko, J. Prilepsky, S. Derevyanko, A. Alvarado, P. Bayvel, and S. Turitsyn, “A lower bound on the per soliton capacity of the nonlinear optical fibre channel,” in *IEEE Information Theory Workshop* (2015), pp. 104–108.
  88. A. Hasegawa and T. Nyu, “Eigenvalue communication,” *J. Lightw. Technol.* **11**(3), 395–399 (1993).
  89. S. Hari, M. Yousefi, and F. Kschischang, “Multi-eigenvalue communication,” *J. Lightw. Technol.* **34**(13), 3110–3117 (2016).
  90. S. T. Le, J. E. Prilepsky, M. Kamalian, P. Rosa, M. Tan, J. D. Ania-Castañón, P. Harper, and S. K. Turitsyn, “Modified nonlinear inverse synthesis for optical links with distributed Raman amplification,” in *European Conference on Optical Communication* (2015), paper Tu.1.1.3.
  91. S. T. Le, I. D. Phillips, J. E. Prilepsky, P. Harper, A. D. Ellis, and S. K. Turitsyn, “Demonstration of nonlinear inverse synthesis transmission over transoceanic distances,” *J. Lightw. Technol.* **34**(10), 2459–2466 (2016).
  92. A. Maruta and Y. Matsuda, “Polarization division multiplexed optical eigenvalue modulation,” in *International Conference on Photonics in Switching* (2015), pp. 265–267.
  93. V. E. Zakharov and A. B. Shabat, “Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media,” *Soviet J. of Exp. and Theo. Phys.* **34**, 62–69 (1972).
  94. S. A. Derevyanko, J. E. Prilepsky, and S. K. Turitsyn, “Capacity estimates for optical transmission based on the nonlinear Fourier transform,” *Nature Commun.* doi: 10.10138/ncomms12710, (2016).
  95. M. I. Yousefi and X. Yangzhang, “Linear and nonlinear frequency multiplexing,” arxiv:1603.04389 (2016).
  96. X. Yangzhang, M. I. Yousefi, A. Alvarado, D. Lavery, and P. Bayvel, “Nonlinear frequency-division multiplexing in the focusing regime,” arxiv:1611.00235 (2016).

## 1. Introduction

Intra-channel and inter-channel fiber nonlinearities are major impairments in coherent transmission systems that limit the achievable transmission distance [1]. Consequently, digital signal processing techniques for compensating or mitigating the effects of fiber nonlinearities and for exploiting fiber nonlinearities have been investigated. Key distinguishing features of these techniques are their complexities and their capabilities to deal with intra-channel and/or inter-channel nonlinearities. An important challenge is to achieve useful improvements in system performance with acceptable levels of computational and implementation complexity.

In broad terms, the techniques for reducing the impact of fiber nonlinearities on system



performance include those that compensate the nonlinearity-induced signal distortion and those that mitigate the distortion by making the signal propagation more tolerant to fiber nonlinearities. They include perturbation solutions to the coupled nonlinear Schrödinger equation (CNLSE), single-channel and multi-channel digital backpropagation, Volterra series nonlinear equalizers, pulse shaping, and advanced modulation formats. Furthermore, a fundamentally different approach exploits fiber nonlinearity by encoding information in the nonlinear Fourier spectrum, thereby raising the prospect of replacing conventional dense wavelength division multiplexing with nonlinear frequency division multiplexing. In this paper, digital signal processing techniques for contending with fiber nonlinearities are reviewed with specific examples illustrating the diversity of techniques that have been explored.

## 2. Perturbation based pre-compensation

The perturbation-based pre-compensation technique is based on approximate time-domain solutions to the CNLSE that express the impact of fiber nonlinearities on a propagating signal as a first-order perturbation term [1, 2]. This approach has been shown to be effective for both pre-compensation [3, 4] and post-compensation [5, 6] of intra-channel fiber nonlinearities. Assuming that the transmitted optical pulses have a Gaussian shape, analytical expressions in terms of the exponential integral function exist for the perturbation expansion coefficients [1, 4]. Extensions of the original approach include an additive-multiplicative model [7], a power weighted model [8–10], and its application to Nyquist pulse shapes [11, 12] and to multi-subcarrier signals, which also serve to mitigate the performance implications of fiber nonlinearities [13, 14].

The perturbation-based technique can be used to pre-compensate accumulated intra-channel fiber nonlinearities with only one computation step for the entire link and can be implemented using one sample per symbol [1, 4]. However, calculation of the nonlinear perturbation involves single and double summations that are functions of the transmitted symbol sequence and perturbation expansion coefficients  $\{C_{m,n}\}$  where  $m$  and  $n$  denote symbol indices relative to the current symbol. Advances aimed at reducing the computational and implementation complexity of this pre-compensation technique include aggressive quantization of the expansion coefficients [15], and the use of symmetric electronic dispersion compensation (SEDC) and root-raised-cosine (RRC) pulse shaping [16]. The quantization of the expansion coefficients has also been considered in the context of simultaneous optimization of the intervals and levels using a minimum mean square error criterion [17] and a decision directed least mean square algorithm [18]. With SEDC, two simplifications result: 1) all the real parts of the coefficients  $\text{Re}[C_{m,n}]$  are zero and 2) all the imaginary parts of the coefficients  $\text{Im}[C_{m,n}]$  are calculated based on half of the link length  $L/2$ . This reduces the dispersion induced pulse spreading and hence the required number of terms in the truncated summations. A RRC pulse shape also reduces the dispersion induced pulse spreading and thus the number of terms in the truncated summations.

The perturbation-based pre-compensation of a signal includes intra-channel self-phase modulation (iSPM), intra-channel cross phase modulation (iXPM) and intra-channel four-wave-mixing (iFWM). With SEDC, the optical field for the current symbol (at time 0) of the x-polarization signal after nonlinear pre-compensation is:

$$A_{0,x}^{\text{out}} = \left( A_{0,x}^{\text{in}} - A_{0,x}^{\text{IFWM}} \right) \cdot \exp(-\Delta\psi_{0,x}) \simeq \left( A_{0,x}^{\text{in}} - A_{0,x}^{\text{IFWM}} \right) \cdot (1 - \Delta\psi_{0,x}), \quad (1)$$

where

$$\Delta\psi_{0,x} = \psi_{0,x} - \mathbb{E}\{\psi_{0,x}\}, \quad (2)$$

$$\psi_{0,x} = 2P \left[ C_{0,0}(L/2) \left( |A_{0,x}|^2 + |A_{0,y}|^2 \right) + \sum_{m \neq 0} C_{m,0}(L/2) \left( 2|A_{m,x}|^2 + |A_{m,y}|^2 \right) \right], \quad (3)$$

$$A_{0,x}^{IFWM} = 2jP^{3/2} \left[ \sum_{m \neq 0, n \neq 0} \text{Im} [C_{m,n}(L/2)] (A_{n,x} A_{m,x} A_{m+n,x}^* + A_{n,y} A_{m,x} A_{m+n,y}^*) + \sum_{m \neq 0} \text{Im} [C_{m,0}(L/2)] (A_{0,y} A_{m,x} A_{m,y}^*) \right]. \quad (4)$$

The corresponding equations for the y-polarization signal are obtained by exchanging the subscripts  $x$  and  $y$  in Eqs. (1) - (4). The nonlinear perturbation coefficients  $\{C_{m,n}\}$  depend on the pulse shape, fiber properties, and fiber length  $L$  [1, 4, 16].  $P$  is the transmitted optical power,  $A_{n,x/y}$  is the sequence of complex transmitted symbols for the x- and y-polarization signals with zero dispersion,  $\mathbb{E}$  denotes expectation, and  $j = \sqrt{-1}$ . Equation (3) represents the phase perturbation due to iSPM and iXPM while Eq. (4) represents the iFWM. It is important to note that for a dual polarization signal there are cross-polarization contributions in Eqs. (3) and (4). The perturbation for the x-polarization signal depends on the transmitted symbol sequences for both the x- and y-polarization signals. The complexity of the algorithm is primarily determined by the second terms in Eq. (3) for iXPM and Eq. (4) for iFWM (and the corresponding equations for the y-polarization signal). The summations are truncated in practice based on the values of  $|C_{m,n}|$  being larger than a specified criterion.

The  $C_{m,n}$  coefficients are fixed for a given transmission spectrum and fiber length. For a RRC pulse shape with a roll-off factor of 0.1 and matched filtering, the coefficients are calculated numerically as an analytical solution is not known [1]

$$C_{m,n}(z) = j\gamma k \left[ \frac{1}{L_{span}} \int_0^{L_{span}} f_{pd}(z') dz' \right] \int_0^z I_{m,n}(z') dz', \quad (5)$$

$$I_{m,n}(z) = \int u_0^*(z, t) u_0(z, t - T_n) u_0(z, t - T_m) u_0^*(z, t - T_{m+n}) dt, \quad (6)$$

where  $\gamma$  is the fiber nonlinear coefficient,  $0 < k \leq 1$  is an optimization factor that may be used to yield the best compensation [11, 18],  $L_{span}$  is the span length,  $f_{pd}(z)$  is the power distribution profile along the link,  $T$  is the symbol period,  $T_m = mT$ ,  $u_0(0, t)$  is the pulse shape with zero accumulated dispersion ( $z = 0$ ), and  $u_0(z, t)$  is the dispersed pulse shape corresponding to a fiber length  $z$  which is calculated according to

$$u_0(z, t) = \mathcal{F}^{-1} \left\{ \mathcal{F}(u_0(0, t)) \exp \left( -j\beta_2 (2\pi f)^2 z/2 \right) \right\}. \quad (7)$$

In Eq. (7),  $\mathcal{F}$  denotes the Fourier transform,  $\mathcal{F}^{-1}$  denotes the inverse Fourier transform,  $f$  is frequency, and  $\beta_2$  is the first order group velocity dispersion coefficient [1].

For a fiber length of 3600 km, with the RRC pulse shape and SEDC,  $|\text{Im}[C_{m,n}(L/2)]|$  is plotted in Fig. 1. The bandwidth of a RRC pulse shape with a roll-off factor of 0.1 yields a small dispersion induced pulse spreading and hence a reduction in the number of terms in truncated approximations to Eqs. (3) and (4) compared to a Gaussian pulse or a RRC pulse with a larger roll-off factor.

For a single 128 Gbit/s polarization-multiplexed (PM) 16QAM signal and transmission over 3600 km of standard single mode fiber with EDFA amplification, the dependence of the bit error ratio (BER) on launch power is shown in Fig. 2(a) for linear post-compensation for dispersion (LC), symmetric linear pre- and post-compensation for dispersion (LC-SEDC), and RRC-SEDC nonlinear pre-compensation. The roll-off factor for the RRC pulse shape was 0.1 and the number of terms in the truncated summations for the RRC-SEDC algorithm was based on  $20 \log_{10} |C_{m,n}/C_{0,0}| > -35$  dB. The dependence of the BER at optimum launch power on fiber length for the three algorithms is shown in Fig. 2(b). For a forward error correction (FEC)

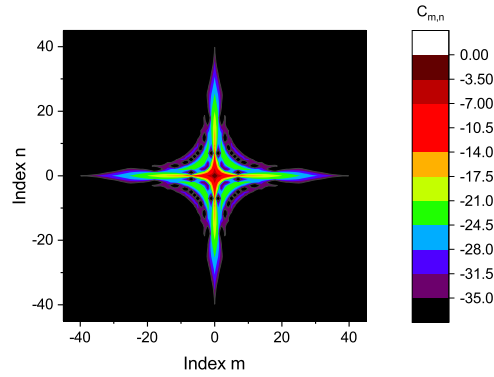


Fig. 1. Example of normalized  $|\text{Im}[C_{m,n}(L/2)]|$  coefficients for 3600 km of standard single-mode fiber with RRC pulse shaping and SEDC.

coding BER threshold of 0.02, transmission over 4200 km of fiber was achieved with RRC-SEDC nonlinear pre-compensation, an increase of 900 km relative to LC and LC-SEDC.

The perturbation-based technique can be used to pre-compensate accumulated intra-channel fiber nonlinearities based on one sample per symbol and one computation step for the entire link. Advances that further reduce the computational and implementation complexity without sacrificing performance would be beneficial. The potential improvements in system performance offered by the technique need to be explored in the context of optical superchannels and flexible-grid networks, including the possibility of extending the algorithm to account for inter-subchannel nonlinearities.

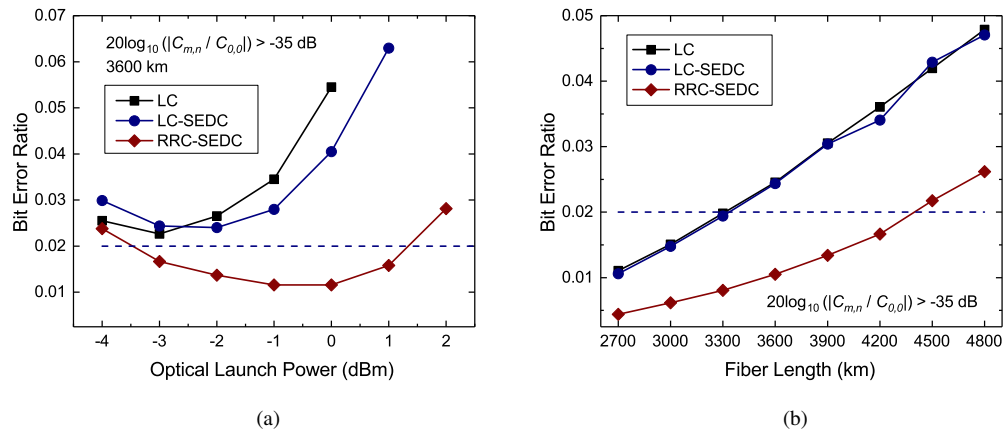


Fig. 2. (a) Dependence of the BER on the optical launch power for a single 128 Gbit/s PM-16QAM signal and a fiber length of 3600 km. (b) Dependence of the BER at optimum launch power on the fiber length. LC: linear post-compensation for dispersion; LC-SEDC: symmetric linear pre- and post-compensation for dispersion; RRC-SEDC: symmetric linear pre- and post-compensation for dispersion, root-raised-cosine pulse shaping, and perturbation-based pre-compensation. Experimental results originally published in [16].



### 3. Wideband digital backpropagation performance

Digital backpropagation (DBP) is arguably the most popular digital signal processing (DSP) technique to compensate for nonlinear optical fiber transmission impairments [19–21]. The effectiveness of the algorithm lies in its ability to fully undo deterministic signal-signal nonlinear interference (NLI) effects.

Despite its theoretical beneficial effects, many factors can contribute to limit the performance of this algorithm, such as: NLI arising from the interaction between the signal and amplified spontaneous emission (ASE) noise [22], polarization-mode dispersion [23–25], DSP complexity at the receiver [24, 26], and limited nonlinearity compensation (NLC) bandwidth. In particular, using analytical tools it has been shown that in fully-loaded wavelength division multiplexing (WDM) systems, DBP gains are severely reduced when DBP is applied over NLC bandwidths that are relatively small compared to the overall transmitted optical bandwidth [27]. If confirmed, this would represent a major setback on the effectiveness of multi-channel DBP performance, as further increasing the NLC bandwidth does not currently appear as a viable option. On the other hand, very few numerical results have been produced to test the accuracy of the available analytical models in predicting the performance of DBP for large NLC bandwidths.

In this section, the analytical tools provided in [28, 29] are validated via numerical results based on the split-step Fourier method (SSFM) in a wideband transmission scenario using multichannel DBP. Then, closed-form expressions are used to describe the behaviour of the signal-to-noise ratio (SNR) gains achievable through DBP.

#### 3.1. Validation of analytical tools for DBP performance estimation

The effect of DBP when applied over a bandwidth  $B_{\text{NLC}}$ , less than or equal to the transmitted bandwidth  $B$ , can be predicted by resorting to a perturbation analysis [30, Sec. II]. To the first-order, the DBP contribution can be considered as a subtraction of a fraction of the received NLI power. Such fraction is equal to the one generated in the forward propagation by the signal within the bandwidth  $B_{\text{NLC}}$  if it was transmitted alone.

The receiver SNR after DBP is applied can be therefore written as

$$\text{SNR} = \frac{P}{N_s P_{\text{ASE}} + [\eta(B, N_s) - \eta(B_{\text{NLC}}, N_s)] P^3 + \eta_{sn} \zeta P_{\text{ASE}} P^2}, \quad (8)$$

where  $P$  is the transmitted power per channel,  $N_s$  is the number of fiber spans,  $P_{\text{ASE}}$  is the ASE noise power over the channel bandwidth,  $\eta(B, N_s)$  is the signal-signal NLI factor over a bandwidth  $B$  and  $N_s$  spans,  $\eta_{sn}$  is the signal-ASE NLI factor over one span,  $B$  is the total transmitted bandwidth,  $B_{\text{NLC}}$  is the NLC bandwidth,  $\zeta = \sum_{k=1}^{N_s} k^{1+\epsilon}$  is the signal-ASE NLI accumulation factor, and  $\epsilon$  is the NLI coherence factor.

In the denominator of Eq. (8), three terms can be distinguished (from left to right): the total accumulated ASE noise power, the residual signal-signal NLI power after DBP is applied, and the signal-ASE NLI power. As discussed in [31, 32], DBP does not modify the signal-ASE NLI power generated in the forward direction. In fact, DBP undoes the signal-ASE NLI originating from the first spans in the forward direction, but replaces it with the one generated by the ASE noise in the last spans in the backward direction.

The  $\eta$  factor and its dependency on system parameters, such as  $B$  and  $N_s$ , vary based on the specific model adopted. For instance, the GN-model [33] offers a simple closed-form expression for  $\eta(B, N_s)$ , although with a certain degree of inaccuracy due to its inability to account for certain features of the transmitted signal, such as the modulation format. More recent models [28, 30, 34] have instead captured the NLI dependence on the modulation format and thus have been shown to be more accurate in the estimation of the NLI power. However, this generally comes at the cost of a higher complexity of the analytical expressions. Recently, in [29], an approximate closed-form expression was proposed for the model in [28], which is derived from an extension

Parameter	Value
<i>Transmission parameters</i>	
Symbol rate	32 Gbaud
Number of channels	31
Channel spacing	33 GHz
Modulation format	PM-16QAM
Pulse shape	Root-raised-cosine
Roll-off factor	0.03
<i>Link parameters</i>	
Fiber attenuation ( $\alpha$ )	0.2 dB/km
Fiber dispersion (D)	17 ps/nm/km
Fiber nonlinearity ( $\gamma$ )	$1.2 (\text{W} \cdot \text{km})^{-1}$
EDFA gain	16 dB
EDFA noise figure	4.5 dB
<i>Simulation parameters</i>	
Number of transmitted symbols	$2^{17}$
Simulation bandwidth	2.04 THz
Numerical method	Adaptive log-step size SSFM [35]

Table 1. Parameters of the system used for the numerical study of DBP performance.

of the GN-model, hence called the enhanced GN-model (EGN). This expression for the analytical estimation of the NLI is used here.

The comparison between analytical and numerical results based on the SSFM is performed for a wideband transmission system, whose parameter values are shown in Table 1. The transmission of  $31 \times 32$  Gbaud PM-16QAM channels with 33 GHz spacing ( $B \approx 1$  THz) is simulated using an adaptive logarithmic step-size SSFM [35]. The transmission link consists of standard single-mode fiber with EDFA amplification. At the receiver, DBP is performed ideally, using the same step-size distribution used in the forward propagation. Ideal polarization demultiplexing is then applied and no carrier phase estimation is used as laser phase noise is neglected.

In Fig. 3, the dependence of the SNR on the transmitted power is shown when either electronic dispersion compensation (EDC) or DBP over different NLC bandwidths is performed at the receiver. It can be observed that the agreement between the analytical expressions and the SSFM simulations is within 0.2 dB for all cases shown. We attribute this residual gap partly to the fact that the closed-form expression for  $\eta(B, N_s)$  strictly holds only for a perfectly rectangular channel spectrum (roll-off factor of 0), whereas the roll-off factor here is set to 0.03. This result confirms the validity of Eq. (8), where  $\eta(B, N_s)$  is obtained from the closed-form expression proposed in [33].

### 3.2. DBP SNR gains

In the previous subsection, the use of closed-form expressions to fully describe DBP performance in a wideband transmission scenario was justified. In this subsection, Eq. (8) is used to describe the analytical behaviour of DBP SNR gain.

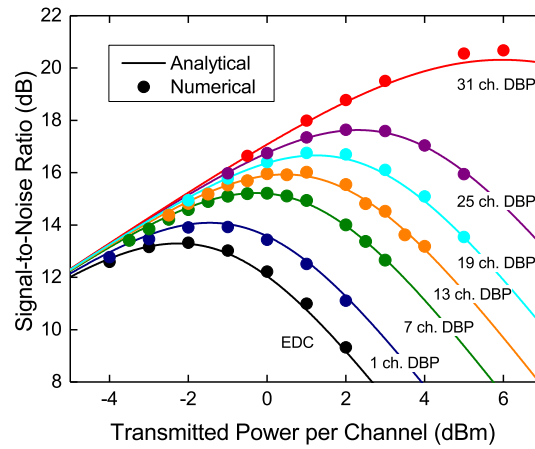


Fig. 3. DBP SNR performance for a transmission of 31×32 Gbaud PM-16QAM channels over 3200 (40×80) km.

For small enough NLC bandwidths, it can be assumed that

$$[\eta(B, N_s) - \eta(B_{\text{NLC}}, N_s)] P^3 \gg \eta_{sn} \zeta P_{\text{ASE}} P^2 \quad (9)$$

and thus the signal-ASE NLI can be neglected in the denominator of Eq. (8). The region where Eq. (9) holds depends on the specific transmission distance and transmitted power. By setting the derivative of Eq. (8) with respect to the transmitted power to zero, the optimum SNR can be found for all NLC bandwidths (including the EDC case). The DBP gain compared to the EDC case (at their respective optimum launch powers) is found in closed-form as

$$G_{\text{DBP}} \approx \sqrt[3]{\frac{\eta(B, N_s)}{\eta(B, N_s) - \eta(B_{\text{NLC}}, N_s)}}. \quad (10)$$

In the regime opposite to the one indicated by Eq. (9), i.e., in a close neighbourhood of the full-field NLC bandwidth, the DBP gain can be approximated as

$$G_{\text{DBP}} \approx \frac{3\sqrt[3]{2}}{4} \left( \frac{2 + \epsilon}{3} \right) \frac{1}{\eta(B, 1)^{1/6} P_{\text{ASE}}^{1/3} N_s^{1/2 + \epsilon/6}}. \quad (11)$$

This approximation holds when Eq. (9) can be considered true in the EDC case, which implies small enough  $\eta_{sn}$ ,  $P_{\text{ASE}}$  and  $N_s$ . However, this is the case for typical transmission scenarios. Two additional assumptions are made in the derivation of Eq. (11): (i) the dependence of  $\eta$  on the number of spans  $N_s$  is assumed for simplicity to be the one predicted by the GN-model, and (ii)  $\eta_{sn} = 3\eta(B, 1)$ , which rigorously holds only when the WDM signal spectrum is flat and its bandwidth  $B$  is equal to the ASE noise bandwidth. The validity of Eq. (11) will be shown in the following.

Eq. (11) shows that the full-field DBP gain is weakly dependent on the ASE noise ( $P_{\text{ASE}}^{-1/3}$ ) and transmitted bandwidth ( $\eta^{-1/6}$ ), whereas it is more strongly dependent on the transmission distance ( $N_s^{-1/2}$ ).

The two asymptotes in Eqs. (10) and (11) are illustrated in Fig. 4(a), where  $G_{\text{DBP}}$  is shown as a function of the NLI reduction factor for different transmission distances. The NLI reduction

factor can be defined as

$$\rho \triangleq \frac{\eta(B, N_s)}{\eta(B, N_s) - \eta(B_{\text{NLC}}, N_s)} \quad (12)$$

and signifies the reduction of signal-signal NLI due to DBP. For small values of NLI reduction, i.e., where signal-signal NLI is dominant compared to signal-ASE NLI, Eq. (10) indicates that the DBP SNR gain increases with a slope of 0.33 dB/dB, i.e., 1 dB higher gain for every 3 dB of suppressed signal-signal NLI. Due to the larger amount of signal-ASE NLI in long-distance transmissions, the gain starts to saturate at smaller values of  $\rho$ . For higher values of  $\rho$ , the gain approaches the full-field gain predicted by Eq. (11).

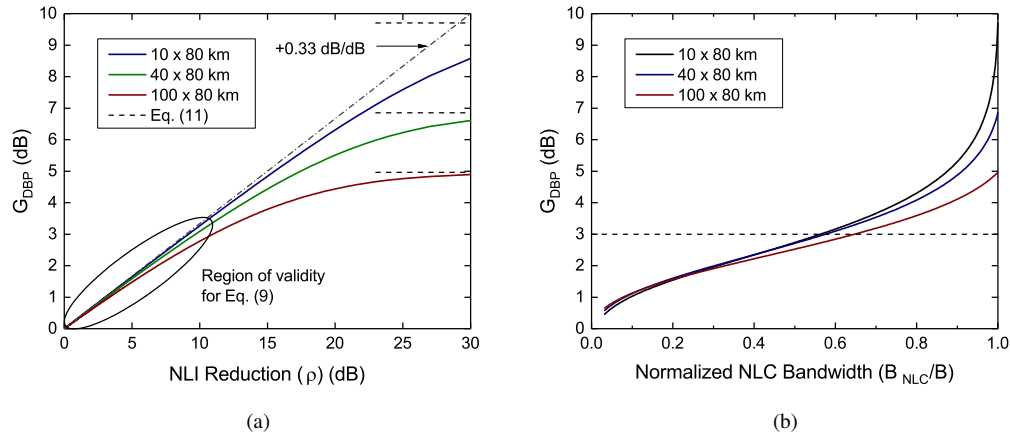


Fig. 4. DBP gain as a function of (a) NLI reduction and (b) normalized NLC bandwidth for the system in Table 1.

Finally, using the closed-form expressions in [29], the DBP gain can be expressed in terms of the NLC bandwidth  $B_{\text{NLC}}$ . This relationship is illustrated in Fig. 4(b), where  $G_{\text{DBP}}$  is shown as a function of  $B_{\text{NLC}}$  normalized with respect to the transmitted bandwidth  $B = 1.023$  THz (see parameters in Table 1), and for different transmission distances. DBP gains are similar (within 0.5 dB difference) for all distances when DBP is applied up to approximately 60% of  $B$ . For small  $B_{\text{NLC}}$  relative to  $B$ , the SNR gain is observed to increase slowly. For instance, in order to achieve 1 dB gain, DBP needs to be applied over approximately 10% of the transmitted bandwidth ( $\approx 100$  GHz), whereas to attain a 3 dB gain, a  $B_{\text{NLC}}$  between 57% ( $\approx 580$  GHz) and 63% ( $\approx 650$  GHz) of  $B$  is required, depending on the transmission distance. A rapid gain increase can instead be obtained when the full-field  $B_{\text{NLC}}$  is approached, particularly for shorter transmission distances. Indeed, in this case, the small amount of residual signal-ASE NLI causes the gain to increase abruptly as the signal-signal NLI is fully cancelled. Higher amounts of signal-ASE NLI instead result in a more gradual increase.

In summary, we have shown, by comparison with SSFM results, that currently available closed-form expressions can accurately predict the receiver SNR of transmission systems employing multichannel DBP to compensate for both intra- and inter-channel NLI. Closed-form relationships between DBP gain and the main system parameters allow quick and intuitive insight into the performance of this algorithm. For NLC bandwidths up to 60% of  $B$ , the relationship between DBP gain and NLI reduction (in dB) is linear through a factor of 1/3. In this region, SNR gains are between 1 and 3 dB. Beyond this region, and as  $B_{\text{NLC}}$  approaches the full-field bandwidth  $B$ , the DBP gain experiences a rapid increase which is dependent on the amount of signal-ASE NLI.

#### 4. Volterra based nonlinear compensation

The Volterra series is a well-known numerical tool for the modelling and compensation of nonlinear dynamic phenomena [36]. It is based on a polynomial expansion, truncated to  $n$ th order, including memory effects through a series of convolution integrals. The Volterra series was first proposed for the modelling of optical fiber transmission systems in [37]. It was applied to solve the NLSE in the frequency-domain, enabling the extraction of a set of  $n$ th order nonlinear transfer functions for a single-mode optical fiber, the so-called Volterra series transfer function (VSTF). The same analytical formulation was also independently developed in [38] in the context of OFDM transmission.

By inverting the 3rd order nonlinear transfer function, an inverse VSTF (IVSTF) was first applied for the compensation of fiber nonlinearities in single-polarization optical transmission [39, 40]. It was shown that, when applied at a low sampling-rate (2 samples per symbol), a 3rd order truncated IVSTF could provide higher performance than split-step-based DBP due to the avoidance of recursive time/frequency transitions [39]. In its polarization multiplexed form, the frequency-domain nonlinear compensated optical field for the x-polarization signal,  $\tilde{A}_x^{\text{NL}}$ , is given by

$$\begin{aligned} \tilde{A}_x^{\text{NL}}(\omega_n, z - L) = & -j\frac{8}{9}\xi\gamma K_1(\omega_n, L) \sum_{m=1}^N \sum_{k=1}^N K_3(\omega_n, \omega_k, \omega_m) \tilde{A}_x(\omega_{n+m-k}, z) \\ & \times [\tilde{A}_x(\omega_k, z) \tilde{A}_x^*(\omega_m, z) + \tilde{A}_y(\omega_k, z) \tilde{A}_y^*(\omega_m, z)], \end{aligned} \quad (13)$$

where  $\tilde{A}_x$  is the frequency-domain received signal in the x-polarization,  $\gamma$  is the nonlinear coefficient,  $L$  is the IVSTF step-size (multiple of the span length,  $L_s$ ),  $0 < \xi \leq 1$  is a free optimization parameter,  $N$  is the fast Fourier transform (FFT) block-size, and  $\omega_n$  is the angular frequency at index  $n$  in the FFT block. The multi-span linear kernel,  $K_1$ , accounts for attenuation and chromatic dispersion as

$$K_1(\omega_n, z) = \exp\left(\frac{\alpha}{2}L_s - j\frac{\beta_2}{2}\omega_n^2 z\right), \quad (14)$$

where  $\alpha$  and  $\beta_2$  are the attenuation and group velocity dispersion coefficients, respectively.  $\beta_2$  is evaluated at the central wavelength of the back-propagated channel. Finally, the multi-span 3rd order nonlinear kernel,  $K_3$ , is given by

$$K_3(\omega_n, \omega_k, \omega_m) = \frac{1 - \exp\left(\alpha L_s - j\beta_2(\omega_k - \omega_n)(\omega_k - \omega_m)L_s\right)}{-\alpha + j\beta_2(\omega_k - \omega_n)(\omega_k - \omega_m)} F(\omega_n, \omega_k, \omega_m), \quad (15)$$

where  $F(\omega_n, \omega_k, \omega_m)$  is the multi-span phased-array factor [38] accounting for the coherent accumulation of nonlinearities between fiber spans

$$F(\omega_n, \omega_k, \omega_m) = \exp\left(-j\frac{\beta_2(\omega_k - \omega_n)(\omega_k - \omega_m)}{2}(L - L_s)\right) \frac{\sin(\beta_2(\omega_k - \omega_n)(\omega_k - \omega_m)L/2)}{\sin(\beta_2(\omega_k - \omega_n)(\omega_k - \omega_m)L_s/2)}. \quad (16)$$

The nonlinear equalized optical field,  $\tilde{A}_x^{\text{NL}}$ , is finally summed with the chromatic dispersion equalization (CDE) signal, yielding the output optical field after each IVSTF step as

$$\tilde{A}_x^{\text{eq}}(\omega_n, z - L) = K_1(\omega_n, L) \tilde{A}_x(\omega_n, z) + \tilde{A}_x^{\text{NL}}(\omega_n, z - L) \quad (17)$$

Note that the equalization of the y-polarization signal is simply obtained by exchanging the subscripts  $x$  and  $y$  in Eqs. (13) and (17).



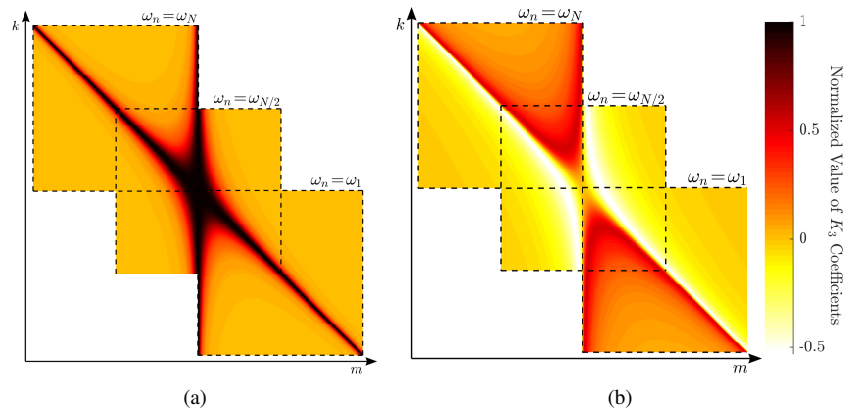


Fig. 5. Normalized (a) real and (b) imaginary components of the 3rd order IVSTF kernel coefficients at three distinct angular frequencies inside a 256-samples FFT block ( $\omega_n = 1$ ,  $\omega_n = 128$  and  $\omega_n = 256$ ). Vertical and horizontal axes correspond to the  $k$  and  $m$  indices in Eq. (13), respectively. The represented IVSTF inverts a single standard single mode fiber span, with signal transmission at 32 Gbaud and sampling rate of 64 GSa/s.

The major challenge associated with the numerical implementation of the IVSTF lies in the  $O(N^2)$  dependence of the total number of operations per equalized sample, arising from the double summation in Eq. (13). This may limit the use of large step-sizes, since the minimum required FFT block length,  $N$ , grows with the accumulated chromatic dispersion. To tackle this issue, several approaches have been addressed. In [41], a simplified IVSTF implementation model with  $O(\log(N))$  complexity was proposed, resorting to parallel nonlinear equalization branches, each of which includes cascaded linear and nonlinear operations in a similar fashion to the SSFM. This approach exploits the linkage between the VSTF and the regular perturbation method [42], employing a frequency-flat approximation to enable time-domain processing of nonlinearities. However, this approximation may affect the performance of the algorithm, which in [41] was shown to underperform relative to single-step per span SSFM-based DBP. Alternatively, in [43], a factorization procedure has been applied to the 3rd order kernel, yielding an  $n$ -steps serial model, similarly enabling a reduction of the complexity down to  $O(\log(N))$ , but also suffering from a performance penalty relative to the full IVSTF model.

Penalty-free approaches have also been proposed, such as the use of symmetric electronic dispersion compensation to reduce the amount of accumulated dispersion to be inverted by the IVSTF [44] and the use of a cascaded IVSTF structure [45], where the position of the linear kernel,  $K_1$ , is changed in order to relax the FFT block length requirements for the evaluation of  $K_3$ .

Another way of reducing the computational effort of the IVSTF is through the inspection and selective pruning of the  $K_3$  coefficients, whose distribution of real and imaginary parts is illustrated in Fig. 5 for an exemplary standard single mode fiber span. For ease of visualization, all coefficients are normalized with respect to the absolute maximum value of the real component. Regular coefficient patterns and column/diagonal symmetries can be clearly observed. Depending on the combination of angular frequencies, different nonlinear phenomena can be identified and categorized as:

- iSPM: when the three optical field components coincide in frequency, i.e., for  $\omega_m = \omega_k = \omega_n$ ;
- iXPM: when the conjugated optical field component coincides in frequency with only one

other component, i.e., for  $\omega_m = \omega_k \neq \omega_n$  or  $\omega_n = \omega_k \neq \omega_m$ ;

- degenerate iFWM: when the two non-conjugated optical field components coincide in frequency, i.e., for  $\omega_k = \omega_{n+m-k}$ ;
- iFWM: for all other possible combinations of  $\omega_m$ ,  $\omega_k$  and  $\omega_n$ .

As can be easily perceived from the inspection of Eq. (15), all iSPM and iXPM occurrences take the same real-valued coefficient, to which corresponds the maximum relative contribution in the  $K_3$  kernel (unitary values in Fig. 5). Based on this inspection of the 3rd order kernel, a simplified Volterra series nonlinear equalizer (VSNE) has been proposed in [46], where the full  $K_3$  matrix is gradually reconstructed as a series of one-dimensional parallel frequency-domain filters, building up from the iSPM+iXPM components and accounting for the symmetries in  $K_3$ . An exact full reconstruction of the  $K_3$  kernel was shown to yield a reduction of the computational complexity by a factor of  $\sim 3$  without any performance penalty [46]. Further simplification can be achieved by exploiting the iXPM-like behavior of the coefficients in the vicinity of the true iXPM components, as can be seen in Fig. 5. Therefore, within a region of validity all coefficients can be forced to the iXPM value incurring only a small error, with a significant reduction in the implementation complexity by avoiding the double summation in (13). This frequency-flat approximation differs from other similar assumptions in the literature [41], since it is associated with an incomplete kernel reconstruction process that departs from the true iXPM component and stops at an optimum number of additional coefficients [46]. Therefore there is a tradeoff between the error generated by the frequency-flat approximation and the error due to an incomplete kernel representation. Building upon this simplified VSNE, equivalent time-domain realizations have also been derived in [48] and experimentally demonstrated in [49], yielding SSFM-like structures with parallel nonlinear compensation branches [50], similar to [41].

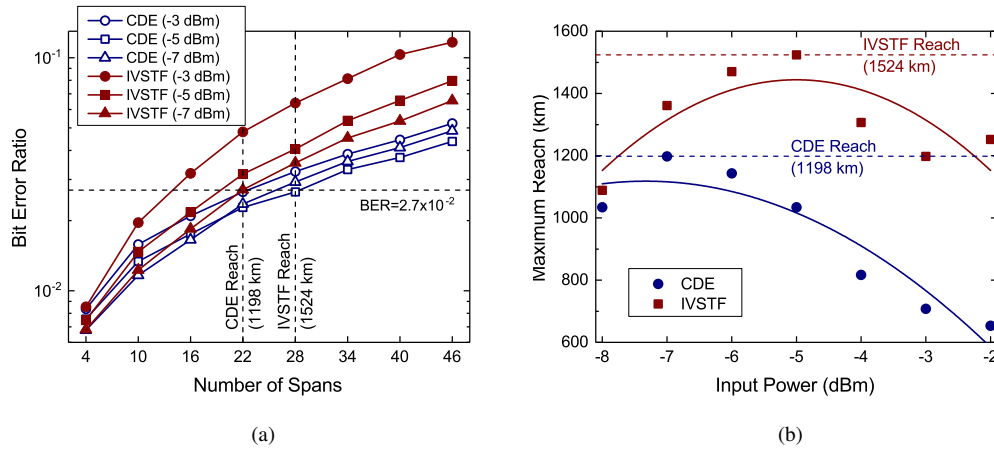


Fig. 6. BER performance and maximum signal reach of 124.8 Gbit/s PM-64QAM enabled by CDE and IVSTF. (a) BER versus number of spans for different channel launch powers; (b) Maximum reach versus launch power. Experimental results originally published in [47].

The IVSTF and its simplified versions proposed in [46] have been experimentally demonstrated in [47], for the nonlinear compensation of a  $10 \times 124.8$  Gbit/s PM-64QAM optical system. The signal was transmitted over pure silica core fiber with an effective area of  $150 \mu\text{m}^2$ , span length of 54.44 km, attenuation of 0.161 dB/km and dispersion parameter of 20.7 ps/nm/km.

The results depicted in Fig. 6 show an improvement of ~25% in the maximum reach (from ~1200 km to ~1500 km) at a BER of  $2.7 \times 10^{-2}$ , provided by nonlinear compensation with the 3rd order IVSTF. A single step IVSTF (step-size  $L$  equal to the full transmission length) was sufficient to achieve the maximum equalization performance. In turn, the frequency-flat simplified VSNE was found to require a total of 4 steps to enable the same maximum reach. Nevertheless, despite the increased processing latency due to 4 cascaded steps, the simplified VSNE was found to reduce the total computational effort by more than 3 orders of magnitude relative to the full matrix-based IVSTF.

Recent advances on IVSTF-based nonlinear compensation have demonstrated similar equalization performance to the widely used SSFM-based DBP, with comparable or even lower computational effort. The full potential of Volterra-based nonlinear compensation is still however far from being achieved. Additional research efforts are required to tackle key implementation aspects such as fast and adaptive coefficient estimation [51] and expansion of the algorithms to account for inter-channel nonlinear compensation in the context of optical superchannels.

## 5. Advanced modulation for nonlinear transmission

The effect of advanced modulation formats on the performance of optical fiber transmission systems can be studied by estimating the achievable information rate (AIR). The AIR provides an upper bound on the maximum data rate, which can be transmitted through a fiber, while also setting a lower bound on the total fiber channel capacity. The AIR is calculated from the mutual information (MI) between the channel input sequence  $X_1^K$  and channel output sequence  $Y_1^K$  of length  $K$

$$I(X; Y) = \lim_{K \rightarrow \infty} \frac{1}{K} I(X_1^K; Y_1^K) = \lim_{K \rightarrow \infty} \frac{1}{K} [\mathcal{H}(X_1^K) - \mathcal{H}(X_1^K | Y_1^K)], \quad (18)$$

where  $\mathcal{H}$  is the entropy function. The AIR is usually expressed in bits/symbol.

The modulation alphabet  $X$  has an effect on the AIR both through the entropy  $\mathcal{H}(X_1^K)$  and the conditional entropy  $\mathcal{H}(X_1^K | Y_1^K)$ . While the former sets an upper bound on the AIR and the spectral efficiency, the latter is a metric of the quality of the received signal, and is usually implicitly used as a design metric. For example, constellation alphabets which reduce nonlinear interference noise (NLIN) increase the signal-to-noise-plus-interference ratio, also referred to as the effective signal-to-noise ratio (SNR) in this section. NLIN is comprised of the signal-signal, signal-ASE and ASE-ASE nonlinear interference effects. This usually leads to reduced uncertainty  $\mathcal{H}(X_1^K | Y_1^K)$ . On the other hand, such constellations can lead to reduced entropy  $\mathcal{H}(X_1^K)$  due to constraints in their construction, leading to a contradiction in the design. It is noted that the output sequence  $y_1^K$  are the samples right before demapping to bits, and the received effective  $\text{SNR} = \mathbb{E}_k[|y_k - x_k|^2]$  thus includes all the penalties from the non-ideal DSP chain (e.g., analog-to-digital conversion, filtering, equalization, phase noise recovery, etc.).

Constellation design in general includes both the positions of the points in the I/Q plane and their probabilities. The former is referred to as *geometric shaping* and the latter as *probabilistic shaping*.

### 5.1. Geometric shaping

One of the first papers on geometric shaping for optical fiber communications was [52]. The main idea was to restrict high-energy symbols in the constellation, thus lowering the peak-to-average power ratio and mitigating the nonlinear effects. To that end, *ring constellations* were studied and optimized for fiber transmission.

A similar approach to constellation design was studied and demonstrated in [53]. Iterative methods were used for optimizing the radii and the number of symbols on each ring with the constraint of 256 symbols in total. An example of the designed *polar modulation* format is given in Fig. 7b, together with the reference 256QAM format in Fig. 7a. The received constellation

diagrams are for a linear AWGN channel with SNR=25 dB and input constellations  $\mathcal{X}$  scaled to unit power. The energy for the polar modulation format is more concentrated towards the origin, thereby allowing for shaping gains over the uniform QAM format in terms of MI for a linear channel. Furthermore, the peak-to-average power ratio is reduced compared to QAM, thus resulting in lower NLIN power. Single channel experimental results were demonstrated for 256 polar modulation [53] with more than 1 dB gain over 256QAM for a 400 km, 28 Gbaud link.

Several other works study geometric signal shaping by imposing constraints on the allowed multi-dimensional sequences, where the considered dimensions are state of polarization and time slots. Lattices were studied in [54] for multi-dimensional constellation design. An optimized minimum Euclidean distance can be (asymptotically) achieved with such constructions, which allows for reduced symbol error rate on a linear AWGN channel. However, a performance penalty was observed in the presence of nonlinearities [54]. Furthermore, bit-to-symbol mapping is non-trivial for such constellations.

*Polarization balanced* multi-dimensional signaling was considered in [55]. Polarization balancing is achieved by constraining the multi-dimensional symbols such that the multi-dimensional energy is constant. Similar to the idea of ring constellations, where the high-energy signals are avoided, such multi-dimensional constellations reduce the NLIN power. The 256 polar modulation from [53] does not change the entropy  $\mathcal{H}(X)$  with respect to 256QAM due to the preserved cardinality of the constellation. In contrast, due to the constellation restriction, this entropy and thereby the spectral efficiency is reduced with multi-dimensional signaling as in [55]. Taking this reduction into account, around 1 dB of net system margin was achieved with an 8D QPSK constellation with respect to the standard BPSK constellation at the same spectral efficiency of 2 bits per time slot in a fully-loaded WDM system with a modulation rate of 35 Gbaud per channel and optical dispersion compensation.

The theoretical gains of such systems were analyzed in [56], where the constellation was restricted to a multi-dimensional ball, for which the mass is concentrated on a multi-dimensional sphere when the number of constellation symbols is large. It was shown that the gains potentially exceed the ultimate shaping gain on an AWGN channel of 1.53 dB. Operating such systems at high spectral efficiency is non-trivial due to the complexity of the DSP at the receiver side. Optimal detection generally requires that each possible input combination of symbols is evaluated, which generally results in an exponential increase in complexity both with the dimensionality (time slots) and the spectral efficiency (cardinality) of the base modulation format (restricted to QPSK in [55]).

## 5.2. Probabilistic shaping

As mentioned, probabilistic shaping attempts to increase the MI by optimizing the probability mass function (PMF)  $p_X(X)$  of the input symbols. This directly results in reduced entropy  $\mathcal{H}(X)$  and thus maximum spectral efficiency of the format. However, near capacity achieving systems operate in a region for which the AIR is not limited by the entropy as much as by the effective SNR at the receiver, thus benefiting from a non-uniform PMF. Probabilistic shaping was performed in [57] by the method of trellis shaping, and near-capacity performance was reported in a simulation. Probabilistic shaping in a 4D space (I/Q dimensions of 16/64QAM in each polarization) was considered in [58], where the 4D PMF was such that, similar to the geometric shaping approach, the points with smaller multi-dimensional amplitude appear more often. Gains of a few hundred kilometers in transmission distance can be achieved with such schemes.

Optimization of the PMF was performed in [59], where the PMF was taken from the Maxwell-Boltzmann (MB) family, for which  $p_X(X = x) \propto \exp(-\lambda|x|^2)$ , i.e., the PMF is also amplitude driven. By carefully optimizing the scaling parameter  $\lambda$ , the PMF can be matched to the channel conditions (the effective SNR). An example of such a PMF for  $\lambda = -0.4$  and a 256QAM constellation is given in Fig. 7(c). Since low-energy points appear more often, the constellation is

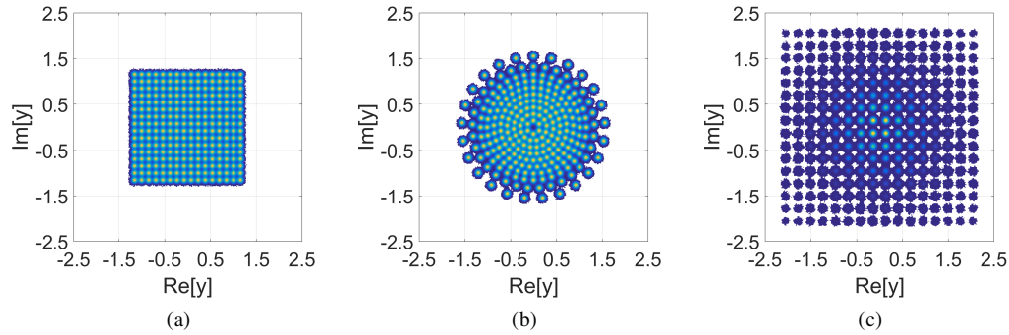


Fig. 7. Constellation diagrams for AWGN channel with SNR=25 dB and constellations, normalized to unit power. (a) Standard 256QAM, (b) geometrically shaped 256 polar modulation, (c) probabilistically shaped 256QAM with Maxwell-Boltzmann distribution. Different probability mass functions (PMFs) result in different scaling, and thereby different Euclidean distance. However, non-uniform PMFs result in reduced entropy  $\mathcal{H}(X)$  and thus reduced maximum spectral efficiency.

scaled, and for unit power and the same SNR as the uniform PMF from Fig. 7(a), the Euclidean distance is increased, resulting in decreased uncertainty  $\mathcal{H}(X_1^K|Y_1^K)$  and increased MI. Gains of up to 400 km in transmission distance were achieved in [59] in a simulation. Experimental demonstration for a selection of MB PMFs was carried out in [60] in combination with a low-density parity check convolutional code. The same gains were experimentally confirmed for a variety of AIRs, which were achieved by rate-matching the independent identically distributed input binary data to the specific MB PMF. Most recently, a system was demonstrated for a transoceanic distance with a record high capacity [61]. The simplicity of the rate matcher, together with its transparency to the FEC makes it attractive for optical fiber communications. An iterative approach to probabilistic shaping was taken in [62], where the PMF was not restricted to the MB family. The PMF was optimized by a modified Blahut-Arimoto algorithm, and it was shown that probabilistic shaping outperforms the geometric shaping scheme from [52]. In order to achieve the non-uniform PMF of the output, a many-to-one bit-to-symbol labeling was proposed in a combination with a convolutional turbo code. It was shown in [63], that this optimization slightly outperforms the MB family, which for two constellation symbols  $x_i$  and  $x_j$  has the restriction of  $p(x_i) > p(x_j)$  for  $|x_i| < |x_j|$ . However, similar experimental gains were achieved as in [60], suggesting that the specific PMF shape is non-consequential in practice under the constraint of independent symbols in each time slot. The performance of the system is given for 256QAM and 1024QAM in Fig. 8 for a 5×10 Gbaud WDM system and distances between 800 and 1700 km at the optimal launch power. The received effective SNR is given in Fig. 8(a). Since the peak-to-average power ratio of the shaped system, particularly for the 1024QAM constellation, is increased, the NLIN noise is enhanced, resulting in slightly decreased effective SNR. However, the AIR with 1024QAM is still superior to the other formats (see Fig. 8(b)) by  $\approx 0.2$  bits/symbol, which translates to 300 km (3 spans) gain at 1200 km ( $\approx 25\%$  reach increase).

It is noted that advanced constellations, such as the ones described here require non-standard equalization and/or phase noise recovery. It was demonstrated in [61, 63], that pilot symbols can be used at a rate of 1-2% for both purposes. This technique also improves the tolerance to phase slips, allows for adaptive equalization, and can potentially be used for frequency and clock recovery. However, improving the DSP performance both in terms of effective received SNR and reduced pilot rate is of interest in practice.

Most of the constellations considered in this section (with the exception of the multi-dimensional



QPSK [55]) operate on a memoryless basis, that is,  $p(x_1^K) = \prod_k p(x_k)$ . Similar gains of about 2-4 fiber spans (200-400 km) are achieved in all the above references under this assumption. In order to improve the gains, PMFs with memory are required. Optimizing such PMFs is not trivial due to the increased dimensionality, and furthermore, optimal processing at the receiver becomes exponentially complex (as mentioned previously) for high spectral efficiency systems. Such multi-dimensional PMFs with jointly optimized geometric and probabilistic constellation properties, and with practical receiver processing are of interest.

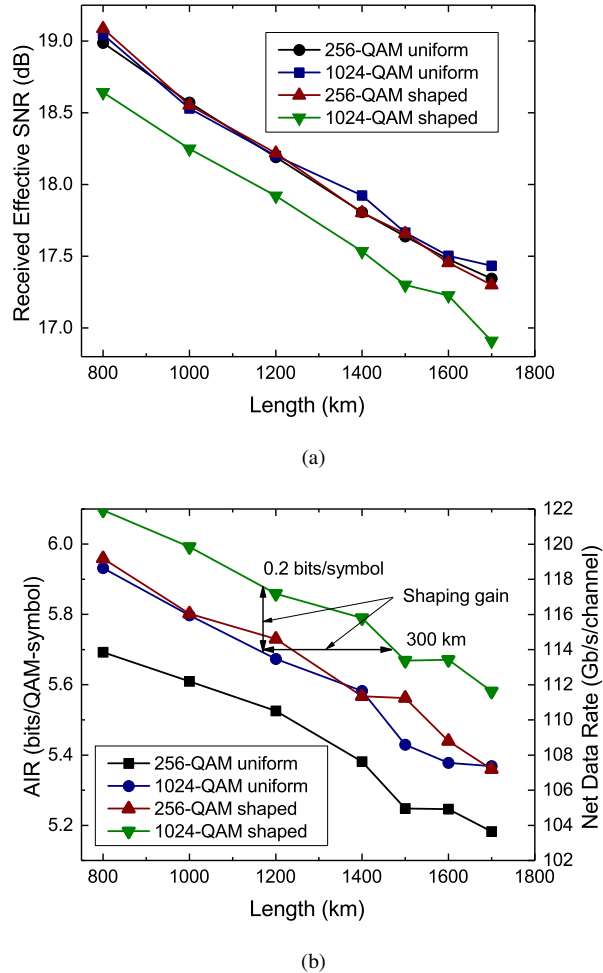


Fig. 8. Performance of probabilistically optimized QAM. (a) Even though 1024QAM with probabilistic shaping results in increased nonlinear distortion and thus reduced effective received SNR, (b) it achieves  $\approx 0.2$  bits/symbol of gain, or equivalently  $\approx 300$  km.

## 6. Encoding in the nonlinear Fourier spectrum

For their discovery in the 1970s of the mathematical framework underlying the nonlinear Fourier transform, C. S. Gardner, J. M. Greene, M. D. Kruskal and R. M. Miura received the prestigious 2006 Leroy P. Steele Prize for a Seminal Contribution to Research, awarded by the American Mathematical Society. In describing this work in [64], the author wrote that “nonlinearity has

undergone a revolution: from a nuisance to be eliminated, to a new tool to be exploited.” This section describes how this tool may be exploited by encoding information in the nonlinear Fourier spectrum (also often called the inverse scattering transform or IST) of a signal transmitted over an optical fiber.

Pulse propagation over an optical link of standard single-mode fiber with ideal distributed Raman amplification is well modelled using the generalized NLSE [65]. In normalized form (see [66]), with time  $t$  and distance  $z$  along the fiber expressed in dimensionless “soliton units”, this equation is given as

$$j \frac{\partial q}{\partial z} = \frac{\partial^2 q}{\partial t^2} - 2s|q|^2 q + n(t, z) \quad (19)$$

where  $j = \sqrt{-1}$ ,  $s \in \{\pm 1\}$ ,  $q(t, z)$  is the complex envelope of the signal, and  $n(t, z)$  is noise, usually modelled as a white Gaussian random process. The first term on the right-hand side expresses the effect on the transmitted waveform of chromatic dispersion, and the second term expresses the effect of Kerr nonlinearity. The equation does not include a loss term, as all losses are assumed to be ideally compensated by Raman amplification. When  $s = -1$ , this equation models signal propagation in the so-called “focusing” regime corresponding to anomalous dispersion (which supports the propagation of soliton pulses), while taking  $s = +1$  gives propagation in the “defocusing” regime corresponding to normal dispersion. In the absence of noise, i.e., with  $n(t, z) = 0$ , Eq. (19) is referred to simply as the NLSE (without the word “generalized”).

In their landmark paper [69], Zakharov and Shabat discovered a Lax pair  $(L, M)$  for the NLSE, thereby establishing its integrability. Fixing  $z$  and writing  $q(t)$  for  $q(t, z)$ , the nonlinear Fourier transform (NFT) of the signal  $q(t)$  is defined in terms of the Zakharov-Shabat system

$$\frac{d\mathbf{v}(t, \lambda)}{dt} = \begin{bmatrix} -j\lambda & q(t) \\ sq^*(t) & j\lambda \end{bmatrix} \mathbf{v}(t, \lambda) \quad (20)$$

where  $\lambda \in \mathbb{C}$  is a spectral parameter—an eigenvalue of the  $L$  operator—and  $\mathbf{v}(t, \lambda)$  is a corresponding  $2 \times 1$  eigenfunction. Let  $u(t, \lambda) = [u_1(t, \lambda), u_2(t, \lambda)]^T$  denote the solution of Eq. (20) under the boundary condition  $\mathbf{v}(t, \lambda) \rightarrow [1 \ 0]^T e^{-j\lambda t}$  as  $t \rightarrow -\infty$ . Define the spectral coefficients  $a(\lambda)$  and  $b(\lambda)$  as

$$a(\lambda) = \lim_{t \rightarrow \infty} u_1(t, \lambda) e^{j\lambda t}, \quad b(\lambda) = \lim_{t \rightarrow \infty} u_2(t, \lambda) e^{-j\lambda t}, \quad (21)$$

and let  $a'(\lambda) = \frac{d}{d\lambda} a(\lambda)$ . Finally, denote the upper-half complex plane (i.e., the set of complex numbers with positive imaginary part) as  $\mathbb{C}^+$ , and let  $\mathbb{D} = \{\lambda \in \mathbb{C}^+ : a(\lambda) = 0\}$ . Since  $a(\lambda)$  is analytic in  $\mathbb{C}^+$ , the set  $\mathbb{D}$  consists of isolated points [66]; furthermore  $\mathbb{D}$  is finite when  $q$  has finite energy. The NFT of  $q(t)$  is the function  $Q : \mathbb{R} \cup \mathbb{D} \rightarrow \mathbb{C}$  defined by

$$Q(\lambda) = \begin{cases} \frac{b(\lambda)}{a(\lambda)}, & \lambda \in \mathbb{R}, \\ \frac{b(\lambda)}{a'(\lambda)}, & \lambda \in \mathbb{D}. \end{cases} \quad (22)$$

Thus, unlike the ordinary Fourier transform, the NFT spectrum generally consists of two components: the continuous spectrum supported on  $\mathbb{R}$  and the discrete spectrum supported on  $\mathbb{D}$ . When  $\mathbb{D}$  is empty, the discrete spectral function is absent. In the defocusing regime (when  $s = +1$ ),  $\mathbb{D}$  is necessarily empty. For small signal amplitudes, the continuous spectrum coincides with the ordinary Fourier transform of  $q(t)$ , and  $\mathbb{D}$  is empty. When present, the discrete spectrum corresponds to the so-called solitonic components of  $q(t)$ . A nonzero signal with a zero continuous spectrum and a discrete spectrum supported on  $N$  points is referred to as an  $N$ -soliton. As noted in [66], the NFT shares many of the properties of the ordinary Fourier transform, including the generalized Parseval identity

$$\int_{t \in \mathbb{R}} |q(t)|^2 dt = \frac{1}{\pi} \int_{\lambda \in \mathbb{R}} \ln(1 + |Q(\lambda)|^2) d\lambda + 4 \sum_{\lambda \in \mathbb{D}} \text{Im}(\lambda). \quad (23)$$

The solitonic signal components influence energy only via the location of the imaginary part of  $\lambda \in \mathbb{D}$ , with larger imaginary part in direct proportion to larger energy. In effect, the NFT is a reformulation of the so-called “scattering data” associated with the IST.

Restoring the  $z$ -dependence,  $Q(\lambda, z)$  denotes the nonlinear Fourier transform of the signal  $q(t, z)$ . The signal  $q(t, 0)$  is applied at the channel input. Under mild assumptions (that  $q(t, 0)$  is absolutely integrable and decays to zero as  $|t| \rightarrow \infty$ ), an extremely simple relationship exists between  $Q(\lambda, 0)$  and  $Q(\lambda, z)$  at any point  $z$ , namely

$$Q(\lambda, z) = Q(\lambda, 0) \exp(4js\lambda^2 z), \quad \lambda \in \mathbb{R} \cup \mathbb{D}. \quad (24)$$

In other words, the NFT of the signal  $q(t, z)$  observed at distance  $z$  is obtained by multiplying the NFT of the input signal  $q(t, 0)$  by a nonlinear frequency response  $H(\lambda, z) = \exp(4js\lambda^2 z)$ . The analogy with linear time-invariant systems is immediate: the NFT plays the same role for systems defined by the NLSE that the ordinary Fourier transform plays for linear time-invariant systems. Note that multiplication by  $H(\lambda, z)$  preserves energy, since for real-valued  $\lambda$ ,  $H(\lambda, z)$  corresponds to an all-pass filter that preserves the energy of the continuous spectral component, while for  $\lambda \in \mathbb{D}$ , multiplication by  $H(\lambda, z)$  does not influence the location of  $\text{Im}(\lambda)$ , which is all that determines the energy of the solitonic component. Energy-preservation is to be expected, since the NLSE models an ideal lossless (and noiseless) system.

An immediate application is an information transmission strategy that is the nonlinear analog of orthogonal frequency-division multiplexing (OFDM), termed nonlinear FDM (or NFDM), that encodes information in the nonlinear spectrum of the signal [68, 70]. Indeed, the idea of encoding information in just the discrete spectrum was first proposed in [88], with recent generalizations given in [73, 89]. A number of recent papers [71–74] have studied various aspects of NFT-based transmission strategies in both the focusing and non-focusing cases. Experimental demonstrations of NFDM schemes and conventional transmission schemes using NFT-based signal detection are described in [75–81]. Numerical methods focused on fast algorithms are described in [82–84].

Of course, actual channels are noisy, and therefore are described by the generalized NLSE Eq. (19). The addition of noise as a forcing term corrupts integrability and the elegant NFT approach does not apply directly. In practice, however, the noise is small, and so can be treated as a perturbation. Depending on the approach taken, various noise models result [68, 85, 86]. Bounds on the “per-soliton” capacity, which include the effects of noise, are provided in [87].

Recent results use numerical methods to estimate the spectral efficiencies that can be achieved using the NFT approach [94–96]. In particular, [94] estimates achievable spectral efficiencies of approximately 10.7 bits per symbol in a 500 GHz bandwidth over a transmission distance of 2000 km in the focusing case ( $s = -1$ ), while [95, 96] estimates achievable rates in excess of 10.5 bits per complex degree-of-freedom at the same distance in both the defocusing case ( $s = +1$ ) and the focusing case. In all three papers, the transmitted information is encoded in the continuous spectrum, fiber parameters are set to practically relevant values, and the transmission power is set to a large value, where the impact of nonlinearity would seriously degrade the performance of conventional transmission techniques. Provided that information is encoded only in the continuous spectrum, there is little difference, from the NFT perspective, between the defocusing and focusing cases, though the latter case does support soliton transmission as well.

Some papers on NFT-based information transmission have incorporated other channel models. It has been shown that the requirement of ideal distributed Raman amplification can be relaxed for modulation of the continuous spectrum [90, 91]. A “lossless path-averaged” (LPA) NLSE was used to deal with lumped amplification from EDFAs as well as non-flat Raman gain profiles. Another recent paper has extended eigenvalue modulation to the polarization multiplexed case [92].

Despite their apparent promise at achieving relatively large spectral efficiencies and enabling transmission at higher launch powers than conventional techniques, considerably more work needs to be done before any of these NFT-based techniques will be competitive in practice. The computational resources required to compute forward and inverse NFTs numerically, even using the fast algorithms of [82–84], is substantial. The impact on the overall transmission system of larger launch powers and nonstandard waveforms needs to be assessed. Interoperability with conventional systems remains a question. Although in principle NFDM does not suffer from deterministic crosstalk effects, achieving nonlinear multiplexing to replace conventional WDM would seem to require access to all co-propagating signals (e.g., the entire C-band), a daunting and presently impractical task. Though more work needs to be done, NFT-based methods may indeed have a role to play in improving future optical fiber transmission systems.

## 7. Summary

Digital signal processing techniques that compensate, mitigate and exploit fiber nonlinearities in coherent optical fiber transmission systems have been reviewed. These include perturbation-based pre-compensation, digital backpropagation, inverse Volterra series transfer function, advanced modulation formats, and encoding in the nonlinear Fourier spectrum.

## Funding

John C. Cartledge and Frank R. Kschischang gratefully acknowledge the support of the Natural Sciences and Engineering Research Council of Canada. Fernando P. Guiomar gratefully acknowledges the support of the European Commission through a Marie Skłodowska-Curie individual fellowship, project Flex-ON (653412). Gabriele Liga gratefully acknowledges the support of the EPSRC through the programme Grant UNLOC (EP/J017582/1). Metodi P. Yankov gratefully acknowledges the support of the Danish National Research Foundation (DNRF123).